DOCUMENT RESUME

ED 407 421 TM 026 443

AUTHOR Love, Angela; And Others
TITLE Examining a Coding Scheme for a Peer Tutoring Study: Agreement, Reliability, or Both?
PUB DATE Oct 96
NOTE 3p.; Paper presented at the Annual Meeting of the Georgia Educational Research Association (Atlanta, GA, October 1996).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Coding; Correlation; *Peer Teaching; *Reliability; *Research Methodology; *Tutoring
IDENTIFIERS *Kappa Coefficient

ABSTRACT

The development of a coding scheme to identify the function of each conversational turn within episodes of conflict in a peer tutoring setting is described, and the scheme, based on Cohen's kappa analysis, is presented. Although 15 codes were developed for the initial effort, 7 codes were finally used to reflect each utterance as: (1) agreement; (2) disagreement; (3) fact; (4) request for information; (5) directive; (6) assertion of solution; and (7) transact. Cohen's kappa, which is a point-by-point analysis of agreement between coders that corrects for change agreement, was used with each dyad. A reliability study was then conducted to evaluate the reliability of each measure, generalizing across coders. Some codes had high reliability; others did not. Combining the study of agreement and reliability was useful in developing the coding scheme. Using Cohen's kappa helped researchers respond to the internal pressure of understanding the measures. Cronbach's intraclass correlation coefficient (Cronbach's alpha) helped researchers respond to the external pressure of conveying to others the accuracy (reliability) of the measures. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 407 421

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
ANGELA LOVE
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

EXAMINING A CODING SCHEME FOR A PEER TUTORING STUDY: AGREEMENT, RELIABILITY, OR BOTH?

Angela Love, Julia L. Stewart, & Ann C. Kruger

TM 026 443

BEST COPY AVAILABLE

Examining A Coding Scheme for A Peer Tutoring Study: Agreement, Reliability, or Both?

Angela Love, Julia L. Stewart, and Ann C. Kruger

Cohen's kappa is often used to evaluate interobserver agreement. It is a point-by-point analysis of agreement between coders that corrects for chance agreement and, because percent agreement does not take into account chance agreement, kappa is a more accurate measure of coder agreement. Bakeman and his colleagues (Bakeman, in press; Bakeman & Gottman, 1986; Bakeman & Quera, 1995) recommend using kappa in evaluating agreement, particularly during training, so that an accurate picture of any basic discrepancies in the coders' perspectives may be revealed.

In addition to interobserver agreement, assessing reliability of the measures used is important in order to generalize across observers within similar populations. Observation studies, however, rarely report reliability of the measures. Bakeman and colleagues recommend a study of reliability using Cronbach's alpha, an intraclass correlation coefficient of reliability, in order to assess if the measure is doing the work it is meant to do. We want to count on the measures being accurate, that is, not being peculiar to any particular observer's perspective.

The purpose of this paper is to examine the development of a coding scheme using Cohen's kappa and evaluate the reliability of the measures using a generalizability coefficient, Cronbach's alpha. The coding scheme identifies the function of each conversational turn within episodes of conflict in a peer tutoring setting. The original coding scheme was based on past and related literature (Kruger, 1993; 1992; Kruger & Tomasello, 1986). The goal, based on these studies, was to characterize conversational turns within episodes of disagreement for each dyad as being transactive or not, that is, identifying whether the use of logical operations on an idea was present or not. In addition, another important characterization of the conflicts was to identify the level of discussion as containing ideas toward solving the task or facts concerning the procedure of the task. In the final analyses, we employed seven codes to reflect each utterance in the conflict: (a) agreement, (b) disagreement, (c) fact, (d) request for information, (e) directive, (f) assertion of solution , and (g) transact.

Initial training, however, was conducted using the original coding scheme of 15 codes all together. Cohen's kappa was run on two dyads, initially, for interobserver agreement for training purposes (kappa = .41). The agreement matrix was revealing in attributing many of the disagreements to confusion between two overarching factors: ideas and facts. Definitions of these two categories were clarified and training continued, as did analyses of each dyad. Kappas ranged from .50 to .80 on seven dyads coded a total of nine times (two tapes were recoded). A close look at the agreement matrices pointed to the need to collapse two sets of codes and to clarify three subcategories of transacts. In our final agreement analysis, kappas ranged from .52 to .92 between coders on individual dyads, with pooled kappa = .73, which meets an acceptable standard of .70 or higher (Bakeman, in press; Bakeman & Gottman, 1986).

Finally, a reliability study was conducted to evaluate the reliability of each measure, generalizing across coders, using the final set of behavioral codes. The results showed respectably high reliability (the range of α = .66 – .98), specifically among the following measures: asserts, agreements, requests for information, and transacts (α = .92, .98, .95, .89, respectively). Two measures were somewhat lower in reliability (disagreements, α = .73 and facts, α = .66) but remained acceptable. The final study of reliability indicates that some of the codes had respectably high reliability and also had a range of variability among subjects. Those codes that had lower reliability could have been due to the lack of variability among subjects within these behaviors coded (Bakeman, in press; Bakeman & Gottman, 1986). It could be argued that these measures are not particularly relevant for this sample or, more likely, that the base rate of these behaviors in this sample under these circumstances is lower such that the amount of sampling in our study was insufficient for higher reliability (Bakeman & Gottman, 1986; Jones, Reid, & Patterson, 1975). It is possible for some codes with high reliability to also have lower agreement revealed by the agreement matrix, because the coders may not have agreed point-by-point but they discriminated equally well among subjects. This was the case with several of our codes.

In conclusion, combining the study of agreement and reliability were both useful in developing a coding scheme that "does the work we want the measure to do" (Bakeman & Gottman, 1986, p. 92). By using Cohen's kappa, we responded to the internal pressure of understanding our measures; by using Cronbach's intraclass correlation coefficient, we responded to the external pressure of conveying to our peers the accuracy of our measures. We recommend to other researchers using coding schemes from behavioral observations to not only consider interobserver agreement using Cohen's kappa during training for coders, but also consider evaluating reliability of the measures using Cronbach's alpha.

Paper presented at the annual meeting of the Georgia Educational Research Association, October, 1996 in Atlanta, Georgia. Requests can be sent to: Angela Love, Department of Educational Psychology & Special Education, Georgia State University, University Plaza, Atlanta, Georgia, 30303. Email address: ALove@gsu.edu.

TM026 443